

TD 1a

Fichiers, formats et extensions

Objectif : distinguer l'extension d'un fichier et son contenu,
sur un même texte enregistré dans plusieurs formats

`cours1/1a_formats/` $\approx 20'$

Deux dossiers : `depart/` et `travail/`

`depart/` contient les fichiers du TD et ne se modifie pas ; `travail/` reçoit vos copies. Une copie abîmée se refait, un original abîmé ne se refait pas.

Attention. Première étape : afficher les extensions, que Windows masque par défaut. Explorateur → Affichage → Afficher → Extensions de noms de fichiers.

	Sur un poste où l'archive est dans Documents
Le fichier de départ	C:\Users\alice\Documents\cours1\la_formats\depart\raven.odt
Sa copie renommée	C:\Users\alice\Documents\cours1\la_formats\travail\raven_odt.pdf

Sous macOS et Linux, `/Users/alice/...` OU `/home/alice/...`, avec des `/`. Les diapositives qui suivent n'écrivent que la fin du chemin, à partir de `cours1/`.

Copier et renommer : à la souris, ou au clavier

Les deux font exactement la même chose. Le clavier va plus vite pour qui le connaît déjà ; il est repris pour lui-même au cours 2.

Dans l'explorateur

1	depart/raven.odt, Ctrl + C
2	dans travail/, Ctrl + V
3	F2, taper raven_odt.pdf, Entrée
4	confirmer le changement d'extension

Dans un terminal (cmd sous Windows)

```
cd cours1\1a_formats
copy depart\raven.odt travail\raven_odt.pdf
copy depart\raven.odt travail\raven.loulou
```

macOS, Linux : `cp depart/raven.odt travail/raven_odt.pdf`

Le terminal est pour ceux qui connaissent : aucune des étapes du TD ne l'exige. Windows demande confirmation quand l'extension change, c'est normal.

Un même document, trois formats

Ouvrir `depart/raven.odt` dans LibreOffice Writer, l'exporter sous deux autres formes dans `travail/`, et comparer ce qu'il reste du texte.

Fichier, dans <code>travail/</code>	Comment
<code>raven.pdf</code>	Fichier → Exporter au format PDF
<code>raven.png</code>	Fichier → Exporter..., type PNG
<code>raven.odt</code>	le fichier de départ, dans <code>depart/</code>

Rouvrir les trois dans LibreOffice : le `.png` s'ouvre dans Draw, comme une image posée sur une page.

Renommer une extension, et voir l'incidence pour leur ouverture

Quatre copies de `depart/raven.odt` dans `travail/`, renommées, puis ouvertes par double-clic.

Nom donné	Ce que le système lance
<code>raven_odt.pdf</code>	un lecteur PDF
<code>raven_odt.jpg</code>	une visionneuse
<code>riri.fifi.loulou.odt</code>	LibreOffice Writer
<code>raven.loulou</code>	rien : il demande quel logiciel choisir

Ouvrir une page html depuis son disque

Double-cliquer sur `depart/raven_brut.html` : le navigateur l'ouvre sans réseau. L'adresse commence par `file:///`.

Fichier ouvert

`raven_brut.html`

`raven_style.html`

`style.css`

Le fichier `.html` est identique dans les deux cas. Seule la ligne `<link rel="stylesheet" href="style.css">` les distingue.

L'adresse que le navigateur affiche

Une adresse web est un chemin de fichier, précédé de la machine où aller le chercher — une **URL**, pour *Uniform Resource Locator*. Celle de la page ouverte depuis le disque a la même forme.

`https://www.ensg.eu/cours/info01/raven.html`
le protocole à quelle machine le chemin sur cette machine le fichier

Dans la barre d'adresse

`file:///C:/Users/alice/Documents/cours1/1a_formats/depart/raven_brut.html`

Pourquoi trois barres après `file:` ?

Un espace dans le nom d'un fichier

Copier `raven_brut.html` dans `travail/` sous le nom `raven brut.html`, l'ouvrir par double-clic, puis regarder l'adresse que le navigateur affiche.

Étape	Ce que vous constatez
copier, puis renommer avec <code>F2</code>	l'explorateur accepte l'espace et l'affiche tel quel
ouvrir par double-clic	
lire l'adresse	
copier l'adresse, la coller dans l'éditeur	

Une adresse ne peut pas contenir d'espace : il y est écrit `%20`, le code du caractère espace. Ce que la barre affiche est une présentation, ce qui se copie est l'adresse.

La table ASCII, caractère par caractère

Les 95 caractères affichables d'ASCII et leur valeur, lue en hexadécimal : le rang donne le premier chiffre, la colonne le second. R est ainsi à la ligne 5_ et dans la colonne 2, soit 52.

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
2_	␣	!	"	#	\$	%	&	'	(	)	*	+	,	-	.	/
3_	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4_	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5_	P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	^	_
6_	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7_	p	q	r	s	t	u	v	w	x	y	z	{		}	~	·

Les valeurs de 00 à 1F ne sont pas des caractères affichables : ce sont des commandes, dont le saut de ligne 0A. 7F non plus.
Les majuscules commencent à 41, les minuscules 20 plus loin, à 61.

Deux éditeurs de texte, les mêmes fichiers

Un éditeur de texte n'affiche rien d'autre que des caractères : il lit chaque octet et montre le caractère correspondant. Le Bloc-notes s'arrête là ; Notepad++ reconnaît l'extension et colore ce qu'il sait lire.

Fichier, dans depart/

raven_une_ligne.txt

style.css

raven_brut.html

raven.odt

Attention. Ne rien enregistrer, et fermer sans sauver : un .odt réenregistré par un éditeur de texte est détruit.